

# Estimation of Online Tool Wear in Turning Processes Using Recurrence Quantification Analysis (RQA)

Srinivasan Radhakrishnan<sup>1</sup>, Yung-Tsun Tina Lee<sup>2</sup> and Sagar Kamarthi<sup>1\*</sup>

**Abstract**—In this work we exploit the underlying dynamics of a turning process captured in force measurements for online flank wear estimation. We transform the sensor signals into feature vectors using recurrence quantification analysis and then estimate flank wear using a gradient boosted regression model. The data is collected by conducting two sets of turning experiments. The first set of data, which has 168 records, is used for training the machine learning model. The second set of data, which has 95 records, is used for testing the performance of the flank wear estimation method. The results indicate that the proposed method gives accurate flank wear estimates. The root mean square of the flank wear estimation for the test data is  $9.87 \times 10^{-5}$  mm.

**Keywords**—Tool wear, flank wear, turning process, recurrence quantification analysis, gradient boosting

## I. INTRODUCTION

Flank wear is a friction related degradation that occurs on the surface of a cutting tool due to the relative motion between the tool and the workpiece. Flank wear increases with time and when the wear progresses beyond a threshold, the tool is either replaced or reworked to ensure optimum performance. Flank wear effects the finishing quality of the workpiece. In the current practice the machine operator replaces the cutting tool before it reaches the end of life or manually examines the tool condition offline by stoppage of the cutting process. In contrast, the sensor-assisted tool monitoring methods examine the tool quality without stopping the turning process. Flank wear estimation involves estimating the width of flank wear by indirect measurements from sensors to assess the state of the system. Online flank wear estimation is one of the main required tasks in automating turning processes. Adequate research efforts have been devoted to develop effective online tool monitoring systems. Most approaches can be classified under physics-based models (analytical models) or machine learning models. Analytical models focus on the underlying laws of physics such as diffusion, adhesion, and abrasion that govern the tool wear in a cutting process. Early, important advancement in analytical models are documented in publications by Koren and Lenz [1], Koren [2], and Usui et al.[3]. On the other hand machine learning methods use direct and indirect measurements of the process state to extract features and build models and then predict a response variable using the model. Historically, multilayer neural networks, adaptive resonance theory networks, and Kohonen's feature maps are

most widely used machine learning models to address tool wear monitoring problems. The application of multilayer neural networks for tool wear monitoring was investigated by Rangwala and Dornfeld [4], Emel [5], Elanayar et al. [6], and Chrystosolouris et al. [7]. Burke [8] studied the performance of adaptive resonance networks for identification of flank wear states (fresh or worn). Kamarthi et al. [9] investigated the application of Kohonen's feature maps for classification of flank wear levels in turning process. In recent years proliferation of data centric approaches, and advent of Internet of Things (IOT) and superior data processing capabilities have enabled wide scale implementation of machine learning models for tool wear monitoring. Appendix A describes the abbreviations used in this paper.

## II. METHOD

An overview of the tool wear estimation method is shown in Figure 1. The force measurements are collected along three directions. Phase synchronization is used to ascertain if the force measurements have captured the dynamics of the process. Once confirmed, one of the force measurements is converted into a *recurrence plot* (RP). From the RP, features are extracted using *recurrence quantification analysis* (RQA). The features are used as an input to a *gradient boosted regression* (GBR) model to estimate tool wear. In the following sections we will discuss in detail the tool wear estimation methodology.

### A. Sensor Data Selection

Most of the methods for tool wear estimation described in literature use cutting force, temperature, vibration, or acoustic emission (AE) measurements, or some combination of these measurements. Force and vibration signals can be measured in three machine-tool coordinate directions: cutting, feed, and radial. Since machining forces in all three coordinate directions are likely to be sensitive to tool wear, though to a varying degree, forces in all three directions are measured in this work. Figure 2 shows the three forces that are generated during a turning operation. Vibration signals in feed direction are more sensitive to flank wear than vibration signals in radial and cutting directions. Studies[10], [11] indicate that vibration in radial and cutting directions exhibit similar characteristics with progressing flank wear. In this work we use force measurements as the input data for flank wear estimation.

In order to determine if the force measurements along all the three machine-tool directions represent the state of the turning process, we use the phase synchronization method.

<sup>1</sup>Mechanical and Industrial Engineering, Northeastern University, Boston, USA [sagar@coe.neu.edu](mailto:sagar@coe.neu.edu)

<sup>2</sup>Systems Integration Division, National Institute of Standards and Technology, Gaithersburg, MD, USA

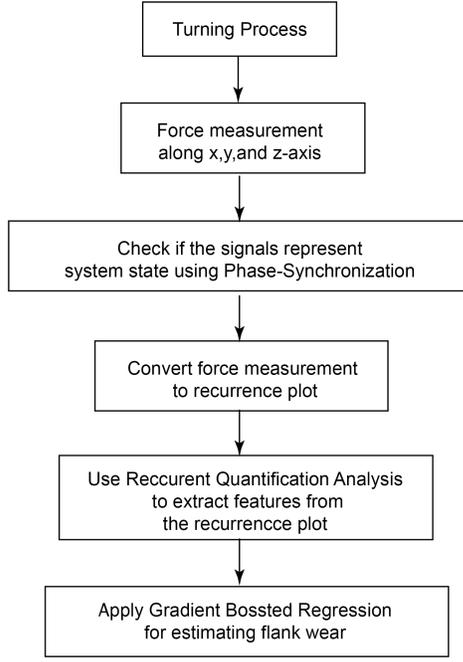


Fig. 1. Overview of the flank wear estimation method

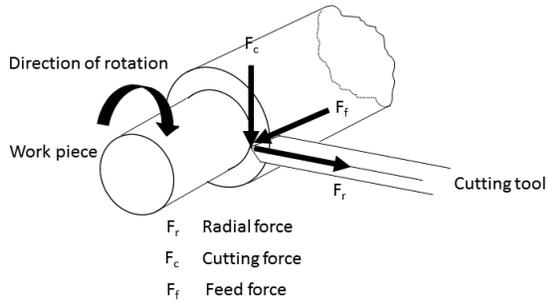


Fig. 2. Force measurements for turning operation

## B. Phase-Synchronization of Sensor Signals and Recurrence plot

A dynamical system like a turning process can be represented in the phase space. The points in the phase space represent the states of the system using Takens algorithm [12], the state of a system at time instance  $i$  can be specified by  $m$  embedded dimensions and  $\tau$  time delays. The dynamics of a time series can be reconstructed in the phase space  $\vec{x}(i)$ :

$$\vec{x}(i) = (x_i, x_{i+2\tau}, \dots, x_{i+(m-1)\tau}) \quad (1)$$

Studies indicate that often  $m = 1$  and  $\tau = 1$  are sufficient to represent the underlying dynamics of the system.

Once a pair of signals (time series) are converted into phase vectors, we use a phase synchronization method to study the underlying dynamics of sensor signals. A set of signals are considered phase synchronized when the difference among their respective phases is bounded [13], the probability  $P(s)$  that the system returns to the  $\epsilon$ -neighborhood of

a former point  $\vec{x}_i$  on the trajectory after  $s$  time interval is computed as [13]:

$$P(s) = \frac{\sum_{i=1}^{N-s} \Theta(\epsilon - \|\vec{x}_i - \vec{x}_{i+s}\|)}{N-s}, \quad (2)$$

where  $s = 1, 2, \dots, K \leq N - 1$ ,  $N$  is the number of measured points  $\vec{x}_i$ ;  $\epsilon$  is a closeness threshold distance;  $\Theta$  is the Heaviside function;  $\|\cdot\|$  is a suitable norm in the phase space under consideration. In this work  $K = 100$  since the time delay does not exceed 100 time steps.

The correlation between  $P_1(s)$  and  $P_2(s)$  for  $s = 1, 2, \dots, K \leq N - 1$  reflects the phase synchronization between the two signals originating from either the same or different systems. It is possible to detect phase synchronization by using the coincidence of the positions of the maxima for the two signals. In this work, the *correlation of probability of recurrence (CPR)*, which is a correlation coefficient, is used as a criterion to quantify the phase synchronization. The *CPR* is defined as follows [13]:

$$CPR = \frac{\sum_{s=1}^K (P_1(s) - \bar{P}_1)(P_2(s) - \bar{P}_2)}{\sqrt{\sum_{t=1}^K (P_1(t) - \bar{P}_1)^2 \sum_{t=1}^K (P_2(t) - \bar{P}_2)^2}} \quad (3)$$

where  $\bar{P}_1$  and  $\bar{P}_2$  are the mean of  $P_1(s)$  and  $P_2(s)$  respectively ( $s = 1, 2, \dots, K$ ). If the two signals are synchronized, the *CPR* is close to positive or negative one; and if they are not synchronized, the *CPR* is close to zero.

We compute *CPR* values between each pair of the three force signals. If the three force signals represent the same dynamical process, then the *CPR* values should be close to 1. In that case we can proceed to use any one of the force measurements for flank wear estimation. We use the time series of one of the force measurements and convert it into a RP to visualize recurrence in phase space. A recurrence plot represents all recurrences in the form of a binary matrix  $R$ , where  $R_{i,j} = 1$  if the state  $x_j$  is in the neighborhood of  $x_i$  in phase space, otherwise  $R_{i,j} = 0$ .  $R_{i,j}$  is computed as follows:

$$R_{i,j}(\epsilon) = \Theta(\epsilon - (\|\vec{x}_i - \vec{x}_j\|)), \quad i, j = 1, \dots, n \quad (4)$$

where  $\vec{x}_i, \vec{x}_j \in \mathfrak{R}^m$  are points in the phase space. Once the recurrence plot is obtained, relevant features are extracted using Recurrence Quantification Analysis (RQA).

## C. Feature Extraction using RQA

Zbilut and Webber [14] introduced RQA for measuring quantitative information hidden in a RP. RQA performs nonlinear data analysis to quantify the number of occurrences and the duration of recurrence of a dynamical system represented by its state space vector. It is actually a measure of information complexity. The main advantage of RQA is that it can provide useful information even for nonlinear and multivariate data. Marwan et al. [15] introduced new measures of complexity to quantify RPs by looking at the small-scale structures such as dots and lines. RPs mostly contains single dots and lines which are vertical/horizontal

or parallel to the mean diagonal referred to as line of identity (LOI). Since a RP matrix is symmetric, horizontal, and vertical lines correspond to each other (the upper right triangle of RP is equal to the lower left triangle of RP); only vertical lines are considered for RQA. The lines capture a typical behavior of the phase space trajectory. While the diagonal lines represent some segments of the phase space trajectory which run parallel for a period of time, the vertical lines represent some segments which remain in the same phase space region for some time duration. The RQA measures are computed based on the recurrence point density, and diagonal and vertical line structures in a RP. The RQA measures were defined by Zbilut and Webber [14] and Marwan et al. [15] and are summarized here for reference:

1) *Recurrence rate (RR)*: *RR* is the density of recurrence points in a RP. This coincides with the correlation sum. It is the estimator of the correlation integral, which is the mean probability that the states at two different times are close. It is computed as:

$$RR = \frac{1}{N^2} \sum_{i,j=1}^N R_{i,j} \quad (5)$$

where  $R_{i,j}$  is an element of the RP and  $N$  is the number of points on the phase space trajectory.

2) *Determinism (DET)*: *DET* is the fraction of recurrence points forming diagonal lines. Diagonal lines represent epochs of similar time evolution of states of the process. *DET* is computed as:

$$DET = \frac{\sum_{l=l_{min}}^N lP(l)}{\sum_{i,j}^N R_{i,j}} \quad (6)$$

where  $P(l)$  is the histogram of the length  $l$  of the diagonal lines,  $l_{min}$  is the minimum acceptable diagonal line length, and  $N$  is the number of points on the phase space trajectory.

3) *Laminarity (LAM)*: *LAM* is the percentage of recurrence points forming vertical lines. Vertical lines indicate intermittency, which is the alternation of phases of apparently periodic and chaotic dynamics. Laminarity is computed as:

$$LAM = \frac{\sum_{v=v_{min}}^N vp(v)}{\sum_{v=1}^N vp(v)} \quad (7)$$

where  $p(v)$  is the histogram of the lengths  $v$  of the vertical lines,  $v_{min}$  is the minimum acceptable vertical line length, and  $N$  is the number of points on the phase space trajectory.

4) *Mean diagonal line length (L)*:  $L$  is the mean prediction time or the inverse of the divergence of the system (K2-entropy). Mean of the diagonal line lengths is computed as follows:

$$L = \frac{\sum_{l=l_{min}}^N lp(l)}{\sum_{l=l_{min}}^N p(l)} \quad (8)$$

where  $p(l)$  is the histogram of the length  $l$  of the diagonal lines,  $l_{min}$  is the minimum acceptable diagonal line length, and  $N$  is the number of points on the phase space trajectory.

5) *Trapping Time (TT)*: *TT* measures the mean time that the system is trapped in a particular state with very slow change. Trapping time is computed as follows:

$$TT = \frac{\sum_{v=v_{min}}^N vp(v)}{\sum_{v=v_{min}}^N p(v)} \quad (9)$$

6) *Longest diagonal line ( $L_{max}$ )*:  $L_{max}$  is the length of the longest diagonal line:

$$L_{max} = \max\{l_i | i = 1, 2, \dots, N_l\} \quad (10)$$

where  $N_l$  is number of diagonal lines in the recurrence plot.

7) *Longest vertical line ( $V_{max}$ )*:  $V_{max}$  is the length of the longest vertical line:

$$V_{max} = \max\{v_i | i = 1, 2, \dots, N_v\} \quad (11)$$

where  $v_l$  is number of vertical lines in the RP.

8) *Entropy (ENTR)*: *ENTR* is the Shannon entropy of the probability distribution of the diagonal lengths. The entropy of the line distribution measures the complexity of the recurrence structure:

$$ENTR = - \sum_{l=l_{min}}^N p(l) \ln(p(l)) \quad (12)$$

where  $p(l)$  is the histogram of the length  $l$  of the diagonal lines,  $l_{min}$  is the minimum acceptable diagonal line length, and  $N$  is the number of points on the phase space trajectory.

The aforementioned RQA measures are used as input features to build the machine learning model.

### III. IMPLEMENTATION OF PROPOSED METHOD

The method shown in Figure 1 is applied to flank wear estimation in a turning process. An extensive set of turning process experiments were conducted to collect data. The details of the workpiece, cutting inserts, sensors, and instrumentation are discussed in the following subsections.

1) *Workpiece and Cutting Tools*: AISI 6150, a chromium-vanadium steel alloy, is chosen as the workpiece material for the experiments. The cutting stock is in the form of cylinders of 914 mm long and 177 mm diameter. The hardness of the workpiece at a radius of 82.55 mm is 425 Bhn and at the center of the cylinder is 360 Bhn. The workpiece is cut with uncoated carbide grade K68 (C2) inserts. K68 is a tough Tungsten-Cobalt (WC-Co) unalloyed grade carbide. The geometric specifications of the insert and the tool holder are SPG-422 and KSBR-164C respectively.

2) *Sensors and Instrumentation*: Experiments are conducted on a 20 HP LeBlond 1610 heavy duty lathe. A three-axis Kistler Z3392/b piezo-electric force dynamometer is used to measure machining forces in cutting, feed, and radial directions. This dynamometer is located underneath the tool post. Force signals are passed through an appropriate set of amplifiers and data acquisition systems to collect the digitized data.

3) *Data Collection*: Cutting is started with a fresh insert edge. Every 60 seconds force measurements are sampled on-line; then flank wear is measured off-line using a toolmaker’s microscope. The data collection at 60 second intervals is carried out until the cutting edge develops an average flank width of at least 0.45 mm, at which point the cutting edge is considered worn out. This limit is chosen according to the criteria recommended by the ISO to define the effective tool life for carbide tools (ISO 3685:1993) [16]. The experiments have shown that an insert edge usually runs for about 10 to 14 minutes before it is considered worn out. The sampling rates for digitization of sensor signals are determined by observing the power spectra of those signals from the preliminary experiments. Force signals are digitized with a sampling frequency of 3 kHz. All the signals are sampled for a length of 4096 points, but only the first 1024 points are considered for computation and analysis purposes.

4) *Experimental Design*: In this work, two sets of experiments, referred to as Set 1 and Set 2 are conducted. In all the experiments the depth of cut is kept constant at 1.27 mm; only the cutting speed and feed are varied. Several preliminary experiments were conducted to decide on the range of operating conditions. It is observed that when the cutting speeds are between 22.86 and 53.34 smpm, there is no formation of build-up edge, no chatter, and no formation of crater wear. Beyond 53.34 smpm, cutting tools start to develop crater wear. When cutting speeds exceeded 83.82 smpm, the tool tip catastrophically begins to deform. Based on these observations, the cutting speeds between 30.48 smpm and 67.06 smpm are used for experimentation.

The Set 1 experiments are conducted according to a full-factorial design with three levels of speed (30.48, 39.62, 48.768 smpm) and five levels of feed (0.16256, 0.22352, 0.28448, 0.34544, 0.39116 mmpr). Accordingly, each one of the 15 ( $= 3 \times 5$ ) tools is used at exactly one combination of speed and feed. The Set 2 experiments are conducted according to a full-factorial design with three levels of speed (30.48, 39.62, 48.768 smpm) and three levels of feed (0.2235, 0.2845, 0.3454 mmpr). So, each one of the 9 ( $= 3 \times 3$ ) tools is used at one combination of speed and feed. It should be noted that the operating conditions used in Set 2 experiments are a subset of the operating conditions used in Set 1 experiments. The operating conditions in both Set 1 and Set 2 experiments cause mostly flank wear on the cutting tools. Set 1 has 168 instances collected over 15 cutting tools and Set 2 has 95 instances collected over 9 cutting tools.

5) *Machine Learning Model*: We investigated Multilayer Neural Network and Gradient Boosted Regression (GBR) as a learning model for estimating flank wear. We found that GBR outperformed multilayer neural networks in key parameters concerned with computation time and estimation error. Hence we used GBR as the machine learning model. Set 1 (training set) is used for training a GBR model and Set 2 is used for validation (test set). The results of the implementation are discussed in the next section.

## IV. RESULTS

Each of the 168 instances from Set 1 generated three force measurements. The force signals were transformed from time domain into phase domain using Equation (1) with  $m = 1$  and  $\tau = 1$ . Phase synchronization defined in Equation (2) was applied to the transformed phase space vectors and the *CPR* values were computed for pairs of force signal components. We found that all the values of *CPR* were close to 1 indicating that any of the three force signals can effectively represent the system state. We use force in main cutting direction in the present work. The cutting force signals in phase space are converted to recurrence plots using Equation (4). A set of features are computed from the RPs using Equations (5-12). We end up with 168 feature vectors that represent the input vectors for the machine learning model. The aforementioned procedure was also followed for the instances from Set 2 which yielded 95 feature vectors. The root mean squared error (RMSE) of Set 2 (test data) estimation is 0.0497 mm. Figure 3 shows the actual (target) and predicted values of flank wear for all nine experiments. Figure 4a shows the progression of actual and predicted flank wear values for the first test tool and Figure 4b shows the box plot of RMSE values for all nine test tools in Set 2.

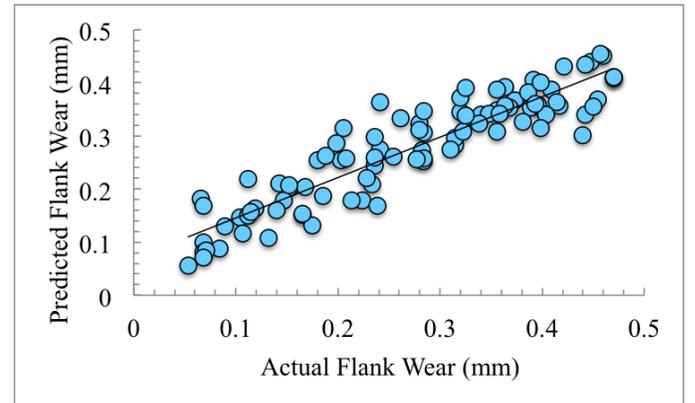


Fig. 3. Actual vs. predicted values for all nine experiments in Set 2 (test data)

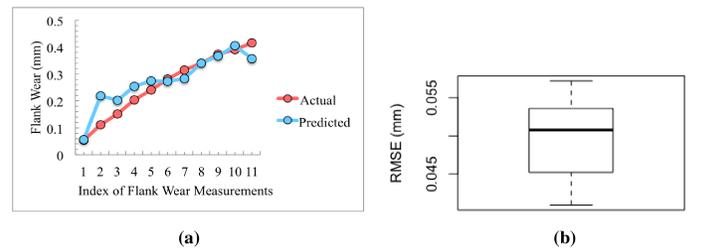


Fig. 4. (a) Progression of actual and predicted flank wear values for the first test tool in Set 2, (b) box plot of root mean squared error for all the nine test tools in Set 2

## V. CONCLUSIONS

The cutting force measurements in a turning process contain information about gradually progressing flank wear.

This information can be successfully extracted and used for online flank wear estimation. The proposed flank wear estimation method provides accurate flank wear estimates within the range of operating conditions that were used during the training data generation. The average RMSE of 9 test tools is 0.0497 mm, which is about 11% deviation from the average flank wear level (0.45 mm) for replacing a cutting tool. This low estimation error makes the proposed estimation method very attractive for real-world applications. In future work we consider the option of using vibration signals and combination of force and vibration signals to further improve the estimation accuracy. In addition we will investigate other machine learning algorithms and feature representation of the measured signals in order to compare the performance of RQA based features with other methods.

## APPENDIX

### A. Abbreviations

- Bhn = Brinell hardness
- CPR = Correlation of probability of recurrence
- GBR = Gradient boosted regression
- mm = Millimeter
- mmpr = Millimeter per revolution
- RMSE = Root mean squared error
- RP = Recurrence plot
- RQA = Recurrence quantification analysis
- smpm = Surface meters per minute

## ACKNOWLEDGMENT

Funding for this research was provided by the National Institute of Standards and Technology under sponsor award number 70NANB15H028.

## DISCLAIMER

Certain commercial software or equipment is identified in this article in order to aid understanding. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

## REFERENCES

- [1] Y. Koren and E. Lenz, "Mathematical model for the flank wear while turning steel with carbide tools," *CIRP Proceedings on Manufacturing Systems*, vol. 1, no. 2, pp. 127–139, 1972.
- [2] Y. Koren, "Flank wear model of cutting tools using control theory," *Journal of Engineering for Industry*, vol. 100, no. 1, pp. 103–109, 1978.
- [3] E. Usui, T. Shirakashi, and T. Kitagawa, "Analytical prediction of cutting tool wear," *Wear*, vol. 100, no. 1-3, pp. 129–151, 1984.
- [4] S. Rangwala and D. Dornfeld, "Sensor integration using neural networks for intelligent tool condition monitoring," *Journal of Engineering for Industry*, vol. 112, no. 3, pp. 219–228, 1990.
- [5] E. Emel, "Tool wear detection by neural network based acoustic emission sensing," *ASME, Dynamic Systems and Control Division Publication*, vol. 28, pp. 79–85, 1991.
- [6] S. Elanayar, Y. C. Shin, and S. Kumara, "Machining condition monitoring for automation using neural networks," in *ASME Winter Annual Meeting, Dallas, TX, USA*, 1990, pp. 85–100.
- [7] G. Chryssolouris, M. Domroese, and P. Beaulieu, "Sensor synthesis for control of manufacturing processes," *Journal of Engineering for Industry(Transactions of the ASME)(USA)*, vol. 114, no. 2, pp. 158–174, 1992.

- [8] L. I. Burke, "An unsupervised neural network approach to tool wear identification," *IIE transactions*, vol. 25, no. 1, pp. 16–25, 1993.
- [9] S. Kamarthi, G. Sankar, P. Cohen, and S. Kumara, "On-line tool wear monitoring using a kohonens feature map," in *Proceeding of the First Artificial Neural Networks in Engineering Conference, St. Louis*, 1991, pp. 639–644.
- [10] S. Rao, "Tool wear monitoring through the dynamics of stable turning," *J. Eng. Ind.(Trans. ASME)*, vol. 108, no. 3, pp. 183–190, 1986.
- [11] A. Sokolowski, J. Liu, and J. Kormol, "On the correlation between the vibration measurement and tool wear in turning. japan," vol. 2, pp. 1075–1081, 1992.
- [12] F. Takens, "Detecting strange attractors in turbulence. in dynamical systems and turbulence, vol 898 of springer lecture notes in mathematics, 366-381," 1981.
- [13] M. C. Romano, M. Thiel, J. Kurths, I. Z. Kiss, and J. Hudson, "Detection of synchronization for non-phase-coherent and non-stationary data," *EPL (Europhysics Letters)*, vol. 71, no. 3, p. 466, 2005.
- [14] J. P. Zbilut and C. L. Webber, "Embeddings and delays as derived from quantification of recurrence plots," *Physics letters A*, vol. 171, no. 3-4, pp. 199–203, 1992.
- [15] N. Marwan, M. C. Romano, M. Thiel, and J. Kurths, "Recurrence plots for the analysis of complex systems," *Physics reports*, vol. 438, no. 5, pp. 237–329, 2007.
- [16] ISO3685:1993, "Tool-life testing with single-point turning tools."